

Your reference: 16113-1300RU1  
Our reference: 2412-136658RU/4193  
Application No.: 2006114696  
Attorney Name: Yury D.Kuznetsov



Page 1 of 8

**TRANSLATION**

**FEDERAL SERVICE ON INDUSTRIAL PROPERTY, PATENTS AND TRADE MARKS (ROSPATENT)**  
bld. 30-1, Berezhkovskaya nab., Moscow, G-59, GSP-5, Russia, 123995 Telephone (8-499) 240-60-15. Fax (8-495) 234-30-58

To No. 2412-136658RU/4193  
(21) Our No. 2006114696/09(015967)

In correspondence, please refer to the application number and report a receipt date of this communication

To "Gorodissky & Partners" Law firm Ltd.,  
B. Spasskaya str., 25, stroenie 3, Moscow, 129090,  
Russia  
Patent attorney Yu. D. Kuznetsov, reg. No. 595

**DECISION ON GRANT OF  
A PATENT FOR INVENTION**

(21) Application No. 2006114696/09(015967)

(22) Application filing date 13.09.2004

As a result of the substantive examination of the present application for invention, it has been stated that

☐ the claimed invention

☒ the claimed group of inventions

relates to the objects of patent rights and complies with the patentability conditions stipulated by the Civil Code of the Russian Federation and, in view of this, a decision to grant a patent for the invention has been taken.

A Report on Examination Results is enclosed.

Head

Signature

B.P. Simonov



Your reference: **16113-1300RU1**  
Our reference: **2412-136658RU/4193**  
Application No.: **2006114696**  
Attorney Name: **Yury D.Kuznetsov**

**Decision on Grant**

Page 2 of 8

**DECISION ON GRANT**  
**PATENT FOR INVENTION**

(21) Application № **2006114696/09(015967)** (22) Date of filing the application **September 13, 2004**  
(24) Date from which industrial property rights may have effect **September 13, 2004**  
(85) Date of commencement of the national phase **May 02, 2006**

**PRIORITY IS FIXED ON DATE**

- ☐ (22) Date of filing the application  
☐ (23) Date of filing of additional materials of to the earlier application №  
☐ (62) ☐ priority date of the application № of from which the present application has been divided up  
☐ filing date of the application № of from which the present application has been divided up  
☐ (66) Filing date of the earlier application №  
☒ (30) Data relating to priority under the Paris Convention  
(31) Number assigned to (32) Date of filing priority (33) Country Claim  
priority application application code  
**10/676,724 September 30, 2003 US**

(86) PCT Application number and date **PCT/US2004/029772** of **September 13, 2004**.

(87) PCT Publication number and date **WO2005/033967** of **April 14, 2005**

(72) Inventor(s) **MITTAL, Vibhu, PONTE, Jay, M., SAHAMI, Mehran, GHEMAWAT, Sanjay, BAUER, John, A., US**

(73) Assignee(s) **GOOGLE INC., US**

(54) Title **SYSTEM AND METHODS FOR SEARCHING USING QUERIES WRITTEN IN A DIFFERENT CHARACTER-SET AND/OR LANGUAGE FROM THE TARGET PAGES**

The Examination department basing on the results of substantive examination of the patent application conducted in respect to

- ☐ originally filed claims ☒ claims amended by the applicant

has revealed their concordance to the requirements of patentability set forth by the Articles 1349 and 1350 of the Civil Code of the Russian Federation and decided to grant the Patent of the Russian Federation for the following claims:

(21) 2006114696/09

(51) IPC

**G06F 17/27** (2006.01)

(57)

1. A method for automatically translating query terms from one language and/or character set to another comprising:  
identifying a first set of anchor text written in a first format and containing a given term;  
identifying a set of documents to which the first set of anchor text points;  
identifying a second set of anchor text written in a second format and pointing to the identified set of documents;  
analyzing the second set of anchor text to determine that a representation of the given term in the first format corresponds to a representation of the given term in the second format.
2. The method of claim 1, in which the first format comprises a first character set, and the second format comprises a second character set.
3. The method of claim 1, in which the first format comprises a first language and the second format comprises a second language.
4. The method of claim 1, in which analyzing the second set of anchor text includes identifying a term that appears most frequently in the second set of anchor text and designating the most frequently appearing term as the representation of the given term in the second format.
5. The method of claim 1, in which analyzing the second set of anchor text comprises:  
calculating a probability that the given term corresponds to a term in the second set of anchor text.
6. The method of claim 5, in which the probability is obtained using at least one of Bayesian methods, histogram smoothing, kernel smoothing, and shrinkage estimators.
7. The method of claim 5, in which the probability that the given term corresponds to a term in the second set of anchor text is obtained by dividing the number of occurrences of the term in the second set of anchor text by the total number of occurrences of all terms in the second set of anchor text.
8. The method of claim 1, in which analyzing the second set of anchor text comprises: calculating a probability that the given term corresponds to each term in the second set of anchor text.
9. The method of claim 1, in which analyzing the second set of anchor text comprises:  
identifying a term that appears most frequently in the second set of anchor text.
10. The method of claim 2, in which the first format is selected from the group consisting of: romaji, romaja, and pinyin; and in which the second character set is selected from the group consisting of: katakana, hiragana, kanji, hangul, hanja, and traditional Chinese characters.

11. The method of claim 1, in which the documents comprise web pages.
12. The method of claim 1, further comprising:  
obtaining a query written in the first format and containing the given term;  
translating the query into the second format based at least in part on said analyzing step;  
searching a database for information written in the second format that is responsive to the translated query.
13. The method of claim 12 in which the steps are performed in the order recited.
14. A method for searching information in one format using queries written in another format comprising:  
obtaining a query written in a first format from a user;  
translating the query into a second format using a probabilistic dictionary, the probabilistic dictionary mapping terms from the first format to the second format;  
searching a database for information that is responsive to the translated query; and  
returning search results written in the second format to the user.
15. The method of claim 14, further comprising:  
obtaining search result selections from the user;  
using said search result selections to modify the probabilistic dictionary of term mappings.
16. The method of claim 15, wherein the modification comprises adjusting at least one probability associated with at least one mapping in the probabilistic dictionary.
17. The method of claim 14, in which the step of translating the query into the second format includes expanding the query.
18. The method of claim 17, in which the expanded query includes alternative encodings of the query terms.
19. The method of claim 17, in which the expanded query includes alternative language translations of the query terms.
20. The method of claim 17, in which the expanded query includes alternative encodings and alternative language translations of the query terms.
21. The method of claim 18, in which the expanded query includes synonyms of the alternative encodings of the query terms.
22. A method for creating a probabilistic dictionary, the probabilistic dictionary mapping terms in a first format to terms in a second format, the method comprising: for a given term, identifying a first set of data in the first format that contains the term;  
identifying a second set of data in the second format that is aligned with the first set of data; and  
analyzing the second set of data to determine one or more probabilities with which the given term maps onto one or more terms in the second set of data.
23. The method of claim 22, further comprising: adding the given term to the dictionary along with one or more probabilities with which the given term maps onto one or more terms in the second set of data.
24. The method of claim 23, further comprising: repeating, for each term to be added to the dictionary, said steps of identifying a first set of data, identifying a second set of data, and analyzing the second set of data.

25. The method of claim 22, in which the first set of data comprises a first set of anchor text pointing to a set of one or more web pages, and in which the second set of data comprises a second set of anchor text pointing to the same set of one or more web pages.
26. The method of claim 22, in which the first set of data comprises a set of text written in a first language, and in which the second set of data comprises the same set of text written in a second language.
27. The method of claim 22, in which the probability with which the given term maps onto a term in the second set of data is calculated by dividing the number of occurrences of the term in the second set of data by the total number of terms in the second set of data.
28. The method of claim 22, further comprising: modifying the probability with which the given term maps onto a term in the second set of data based, at least in part, on an analysis of a user's selection of search results.
29. The method of claim 22, further comprising: modifying the probability with which the given term maps onto a term in the second set of data based, at least in part, on an analysis of a user's previous queries.
30. A computer-readable medium, comprising instructions, which when executed by a computer system, are operable to cause the computer system to perform acts comprising:  
identifying a first set of anchor text written in a first format and containing a given term;  
identifying a set of web pages to which the first set of anchor text points;  
identifying a second set of anchor text written in a second format and pointing to the identified set of web pages;  
determining a probability that a representation of the given term in the first format corresponds to a representation of the given term in the second format.
31. The computer-readable medium of claim 30, further including instructions, which when executed by the computer system, are operable to cause the computer system to perform acts comprising:  
modifying the probability that a representation of the given term in the first format corresponds to a representation of the given term in the second format based, at least in part, on an analysis of a user's selection of search results.
32. The computer-readable medium of claim 30, further including instructions, which when executed by the computer system, are operable to cause the computer system to perform acts comprising:  
modifying the probability that a representation of the given term in the first format corresponds to a representation of the given term in the second format based, at least in part, on an analysis of a user's previous queries.
33. The computer-readable medium of claim 30, in which the probability is determined using at least in part one of Bayesian methods, histogram smoothing, kernel smoothing, and shrinkage estimators.
34. A method for translating a query provided in a first language or character set to a second language or character set comprising:  
identifying a first body of text written in a first format;  
identifying a second body of text written in a second format, the second body of text being aligned with the first body of text;  
creating a dictionary of translations between terms in the first body of text and terms in the second body of text by comparing the occurrence of terms in the first body of text with the occurrence of terms in the second body of text.

35. A method of claim 34, in which the dictionary of translations includes one or more probabilities associated with the translations.
36. A method of claim 34, in which the first format comprises a first character set and the second format comprises a second character set.
37. A method of claim 34, in which the first format comprises a first language and the second format comprises a second language.
38. A method of claim 34, in which the first body of text comprises anchor text and the second body of text comprises anchor text.
39. A method for performing searches using potentially ambiguous queries comprising:  
receiving a query containing at least one query term written in a first format;  
translating the query term into a plurality of variants written in a second format; and  
using one or more of the variants to search for information written in the second format that is responsive to the query.
40. The method of claim 39, in which the first format comprises a sequence of numbers entered from a telephone keypad; and in which the second format comprises alphanumeric text.
41. The method of claim 39, further comprising:  
obtaining the one or more variants by discarding variants in the plurality of variants that are not part of a predefined lexicon.
42. The method of claim 39, further comprising:  
obtaining the one or more variants by discarding variants in the plurality of variants that contain predefined low-probability character combinations.
43. The method of claim 39, in which the first format comprises alphanumeric text written in a character set selected from the group consisting of romaji, romaja, and pinyin; and in which the second format comprises alphanumeric text written in a character set selected from the group consisting of kanji, katakana, hiragana, hangul, hanja, and traditional Chinese characters.
44. A method for performing searches using potentially ambiguous queries comprising:  
receiving a numeric query entered from a telephone keypad;  
translating the numeric query into a group of potential alphanumeric translations in a first format;  
discarding potential translations that are determined to include predefined lowprobability character combinations;  
translating the remaining alphanumeric translations from the first format to a second format using a probabilistic dictionary; and  
performing a search using the alphanumeric translations in the second format.
45. The method of claim 44, in which the first format comprises text written in a character set selected from the group consisting of romaji, romaja, and pinyin; and in which the second format comprises text written in a character set selected from the group consisting of kanji, katakana, hiragana, hangul, hanja, and traditional Chinese characters.

Your reference: **16113-1300RU1**  
Our reference: **2412-136658RU/4193**  
Application No.: **2006114696**  
Attorney Name: **Yury D.Kuznetsov**

**Decision on Grant**

Page 7 of 8

(56) RU 2039376 C1, 09.07.1995

KR 20030044949 A, 09.06.2003

WO 0075808 A1, 14.12.2000

EP 0597611 A2, 18.05.1994

WO 03042870 A1, 22.05.2003

In publishing information on issue of the patent the specification shall be used as original editorial materials submitted by the Applicant.

In publishing information on issue of the patent the drawings will be used as originally filed.

Enclosure: Abstract performed by Examination on 1 sheet in 1 copy.

Your reference: 16113-1300RU1  
Our reference: 2412-136658RU/4193  
Application No.: 2006114696  
Attorney Name: Yury D.Kuznetsov

Decision on Grant

Page 8 of 8

Publish with fig. № 7

Enclosure

To the application №2006114696/09

(54) SYSTEM AND METHODS FOR SEARCHING USING QUERIES WRITTEN IN A DIFFERENT CHARACTER-SET AND/OR LANGUAGE FROM THE TARGET PAGES

#### ABSTRACT

(57) The present invention relates to information search and retrieval. A technical result is performing searches using queries that are written in a character set or language that is different from the character set or language of the documents that are to be searched and providing relevant search results. For this purpose is received a sequence of ambiguous information components from a user and translated its into one or more corresponding sequences of less ambiguous information components. These sequences of less ambiguous information are provided as an input data to a search engine. The search results are obtained from the search engine and are presented to the user. A translation between these character sets and/or languages can be performed by examining the use of terms in aligned text. Probabilities can be associated with each possible translation. Refinements can be made to these probabilities by examining user interactions with the search results. 7 in. cl., 38 dep. cl.; 16 drawings.



**ФЕДЕРАЛЬНАЯ СЛУЖБА ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,  
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ  
(РОСПАТЕНТ)**

Бережковская наб., 30, корп. 1, Москва, Г-59, ГСП-5, 123995. Телефон (8-499) 240- 60- 15. Факс (8-495) 234- 30- 58

На № 2412-136658RU/4102 от 18.11.2008

Наш № 2006114696/09(015967)

*При переписке просьба ссылаться на номер заявки  
и сообщать дату получения данной корреспонденции*

129090, Москва, ул. Б.Спаская, 25, стр.3,  
ООО "Юридическая фирма Городисский и  
Партнеры", пат.пов. Ю.Д.Кузнецову, рег.№  
595

20 ЯНВ 2009

Date G&P: 23/01/2009



0004790808

**РЕШЕНИЕ  
о выдаче патента на изобретение**

(21) Заявка № 2006114696/09(015967)

(22) Дата подачи заявки 13.09.2004

В результате экспертизы заявки на изобретение по существу установлено, что

☐ заявленное изобретение

☒ заявленная группа изобретений

относится к объектам патентных прав и соответствует условиям патентоспособности, предусмотренным Гражданским кодексом Российской Федерации, в связи с чем принято решение о выдаче патента на изобретение:

Заключение по результатам экспертизы прилагается.

Приложение: на 10 л. в 1 экз.

Руководитель

Б.П.Симонов

(21) 2006114696/09

(51) МПК

**G06F 17/27(2006.01)**

(57)

1. Способ автоматического перевода терминов запроса из одного языка и/или набора символов в другой, содержащий этапы, на которых:  
идентифицируют первое множество текста привязки, написанного в первом формате и содержащего данный термин;  
идентифицируют множество документов, на которые указывает первое множество текста привязки;  
идентифицируют второе множество текста привязки, написанного во втором формате и указывающего на идентифицированное множество документов;  
анализируют второе множество текста привязки, чтобы определить, что представление данного термина в первом формате соответствует представлению данного термина во втором формате.
2. Способ по п. 1, в котором первый формат содержит первый набор символов, а второй формат содержит второй набор символов.
3. Способ по п. 1, в котором первый формат содержит первый язык, а второй формат содержит второй язык.
4. Способ по п. 1, в котором анализ второго множества текста привязки включает в себя идентификацию термина, который появляется наиболее часто во втором множестве текста привязки, и обозначение наиболее часто появляющегося термина как представления данного термина во втором формате.
5. Способ по п. 1, в котором анализ второго множества текста привязки содержит этап, на котором:

вычисляют вероятность того, что данный термин соответствует термину во втором множестве текста привязки.

6. Способ по п. 5, в котором вероятность получается с использованием, по меньшей мере, одного из байесовских методов, сглаживания гистограммы, сглаживания функции влияния и оценок сокращения.

7. Способ по п. 5, в котором вероятность того, что данный термин соответствует термину во втором множестве текста привязки, получается путем деления количества вхождений термина во второе множество текста привязки на общее количество вхождений всех терминов во втором множестве текста привязки.

8. Способ по п. 1, в котором анализ второго множества текста привязки содержит этап, на котором:

вычисляют вероятность того, что данный термин соответствует каждому термину во втором множестве текста привязки.

9. Способ по п. 1, в котором анализ второго множества текста привязки содержит этап, на котором:

идентифицируют термин, который появляется наиболее часто во втором множестве текста привязки.

10. Способ по п. 2, в котором первый формат выбирается из группы, состоящей из: ромадзи, ромадза и пиньинь; и в котором второй набор символов выбирается из группы, состоящей из: кандзи, катакана, хирагана, хангиль, хандза и традиционных китайских символов.

11. Способ по п. 1, в котором документы содержат веб-страницы.

12. Способ по п. 1, дополнительно содержащий этапы, на которых: получают запрос, написанный в первом формате и содержащий данный термин;

переводят запрос во второй формат, по меньшей мере, частично на основе упомянутого этапа анализа;

ищут в базе данных информацию, написанную во втором формате, которая соответствует переведенному запросу.

13. Способ по п. 12, в котором этапы выполняются в перечисленном порядке.

14. Способ поиска информации в одном формате с использованием запросов, записанных в другом формате, содержащий этапы, на которых: получают запрос от пользователя, написанный в первом формате; переводят запрос во второй формат, используя вероятностный словарь, при этом вероятностный словарь отображает термины из первого формата во второй формат; ищут в базе данных информацию, которая соответствует переведенному запросу; и возвращают пользователю результаты поиска, написанные во втором формате.
15. Способ по п. 14, дополнительно содержащий этапы, на которых: получают от пользователя варианты выбора результатов поиска; используют упомянутые варианты выбора результатов поиска для модифицирования вероятностного словаря отображений терминов.
16. Способ по п. 15, в котором модификация содержит корректировку, по меньшей мере, одной вероятности, ассоциативно связанной, по меньшей мере, с одним отображением в вероятностном словаре.
17. Способ по п. 14, в котором этап перевода запроса во второй формат включает в себя расширение запроса.
18. Способ по п. 17, в котором расширенный запрос включает в себя альтернативные кодировки терминов запроса.
19. Способ по п. 17, в котором расширенный запрос включает в себя альтернативные языковые переводы терминов запроса.
20. Способ по п. 17, в котором расширенный запрос включает в себя альтернативные кодировки и альтернативные языковые переводы терминов запроса.
21. Способ по п. 18, в котором расширенный запрос включает в себя синонимы альтернативных кодировок терминов запроса.
22. Способ для создания вероятностного словаря, причем упомянутый вероятностный словарь отображает термины в первом формате в термины во втором формате, упомянутый способ содержит этапы, на которых: для

данного термина идентифицируют первое множество данных в первом формате, которое содержит термин;

идентифицируют второе множество данных во втором формате, которое выровнено с первым множеством данных; и

анализируют второе множество данных, чтобы определить одну или более вероятностей, с которыми данный термин отображается в один или более терминов во втором множестве данных.

23. Способ по п. 22, дополнительно содержащий этап, на котором: добавляют данный термин в словарь вместе с одной или более вероятностями, с которыми данный термин отображается в один или более терминов во втором множестве данных.

24. Способ по п. 23, дополнительно содержащий этап, на котором: повторяют для каждого термина, который нужно добавить в словарь, упомянутые этапы идентификации первого множества данных, идентификации второго множества данных и анализа второго множества данных.

25. Способ по п. 22, в котором первое множество данных содержит первое множество текста привязки, указывающего на множество из одной или более веб-страниц, и в котором второе множество данных содержит второе множество текста привязки, указывающего на то же множество из одной или более веб-страниц.

26. Способ по п. 22, в котором первое множество данных содержит множество текста, написанного на первом языке, и в котором второе множество данных содержит то же множество текста, написанного на втором языке.

27. Способ по п. 22, в котором вероятность, с которой данный термин отображается в термин во втором множестве данных, вычисляется посредством деления количества вхождений термина во втором множестве данных на общее количество терминов во втором множестве данных.

28. Способ по п. 22, дополнительно содержащий этап, на котором: модифицируют вероятность, с которой данный термин отображается в

термин во втором множестве данных, по меньшей мере, частично на основе анализа пользовательского выбора результатов поиска.

29. Способ по п. 22, дополнительно содержащий этап, на котором: модифицируют вероятность, с которой данный термин отображается в термин во втором множестве данных, по меньшей мере, частично на основе анализа предыдущих пользовательских запросов.

30. Машиночитаемый носитель, содержащий команды, которые, будучи исполненными вычислительной системой, работают, чтобы заставить вычислительную систему выполнить действия, содержащие: идентификацию первого множества текста привязки, написанного в первом формате и содержащего данный термин;

идентификацию множества веб-страниц, на которые указывает первое множество текста привязки;

идентификацию второго множества текста привязки, написанного во втором формате и указывающего на идентифицированное множество веб-страниц;

определение вероятности того, что представление данного термина в первом формате соответствует представлению данного термина во втором формате.

31. Машиночитаемый носитель по п. 30, дополнительно включающий в себя команды, которые, будучи исполненными вычислительной системой, работают, чтобы заставить вычислительную систему выполнить действия, содержащие:

модификацию вероятности того, что представление данного термина в первом формате соответствует представлению данного термина во втором формате, по меньшей мере, частично на основе анализа пользовательского выбора результатов поиска.

32. Машиночитаемый носитель по п. 30, дополнительно включающий в себя команды, которые, будучи исполненными вычислительной системой, работают, чтобы заставить вычислительную систему выполнить действия, содержащие:

модификацию вероятности того, что представление данного термина в первом формате соответствует представлению данного термина во втором формате, по меньшей мере, частично на основе анализа предыдущих пользовательских запросов.

33. Машиночитаемый носитель по п. 30, в котором вероятность определяется с использованием, по меньшей мере, частично одного из байесовских методов, сглаживания гистограммы, сглаживания функции влияния и оценок сокращения.

34. Способ перевода представленного на первом языке или наборе символов запроса на второй язык : или набор символов, содержащий этапы, на которых:

идентифицируют первую часть текста, написанную в первом формате;

идентифицируют вторую часть текста, написанную во втором формате, вторую часть текста, выравниваемую с первой частью текста;

создают словарь переводов между терминами в первой части текста и терминами во второй части текста посредством сравнения вхождения терминов в первой части текста с вхождением терминов во второй части текста.

35. Способ по п. 34, в котором словарь переводов включает в себя одну или более вероятностей, ассоциативно связанных с переводами.

36. Способ по п. 34, в котором первый формат содержит первый набор символов, а второй формат содержит второй набор символов.

37. Способ по п. 34, в котором первый формат содержит первый язык, а второй формат содержит второй язык.

38. Способ по п. 34, в котором первая часть текста содержит текст привязки, и вторая часть текста содержит текст привязки.

39. Способ для выполнения поисков с использованием потенциально неоднозначных запросов, содержащий этапы, на которых:

принимают запрос, содержащий, по меньшей мере, один термин запроса, написанный в первом формате;

переводят термин запроса во множество вариантов, написанных во втором формате; и

используют один или более вариантов для поиска информации, написанной во втором формате, которая является ответной на запрос.

40. Способ по п. 39, в котором первый формат содержит последовательность цифр, введенных с клавишной панели телефона; и в котором второй формат содержит буквенно-цифровой текст.

41. Способ по п. 39, дополнительно содержащий этап, на котором: получают один или более вариантов посредством отбрасывания вариантов в множестве вариантов, которые не являются частью предопределенного лексикона.

42. Способ по п. 39, дополнительно содержащий этап, на котором: получают один или более вариантов посредством отбрасывания вариантов в множестве вариантов, которые содержат предопределенные маловероятные сочетания символов.

43. Способ по п. 39, в котором первый формат содержит буквенно-цифровой текст, написанный в наборе символов, выбранном из группы, состоящей из ромадзи, ромадза и пиньинь; и в котором второй формат содержит буквенно-цифровой текст, написанный в наборе символов, выбранном из группы, состоящей из кандзи, катакана, хирагана, хангиль, хандза и традиционных китайских символов.

44. Способ для выполнения поисков с использованием потенциально неоднозначных запросов, содержащий этапы, на которых:

принимают цифровой запрос, введенный с клавишной панели телефона; переводят цифровой запрос в группу потенциальных буквенно-цифровых переводов в первом формате;

отбрасывают потенциальные переводы, которые определяются как включающие <sup>предопределенные</sup> маловероятные сочетания символов;

переводят оставшиеся буквенно-цифровые переводы из первого формата во второй формат, используя вероятностный словарь; и выполняют поиск, используя буквенно-цифровые переводы во втором формате.

45. Способ по п. 44, в котором первый формат содержит текст, написанный в наборе символов, выбранном из группы, состоящей из ромадзи, ромадза



и пиньинь; и в котором второй формат содержит текст, написанный в наборе символов, выбранном из группы, состоящей из кандзи, катакана, хирагана, хангиль, хандза и традиционных китайских символов.

(56) RU 2039376 C1, 09.07.1995  
KR 20030044949 A, 09.06.2003  
WO 0075808 A1, 14.12.2000  
EP 0597611 A2, 18.05.1994  
WO 03042870 A1, 22.05.2003

При публикации сведений о выдаче патента будет использовано описание в первоначальной редакции заявителя.

При публикации сведений о выдаче патента будут использованы первоначальные чертежи.

Приложение: Реферат, скорректированный экспертизой, на 1 л. в 1 экз.

Заместитель заведующего отделом  
электрорадиотехники ФГУ ФИПС



Н.Н.Маркелова

Н.С.Лаврентьева  
8 (495) 730 76 51

К заявке № 2006114696/09

**(54) СИСТЕМЫ И СПОСОБЫ ДЛЯ ПОИСКА С ИСПОЛЬЗОВАНИЕМ ЗАПРОСОВ, НАПИСАННЫХ НА ЯЗЫКЕ И/ИЛИ НАБОРЕ СИМВОЛОВ ОТЛИЧНОМ ОТ ТАКОВОГО ДЛЯ ЦЕЛЕВЫХ СТРАНИЦ**

**Реферат**

(57) Изобретение относится к поиску и выборке информации. Техническим результатом является обеспечение возможности выполнения поиска с использованием запросов, написанных в наборе символов или языке, который отличается от набора символов или языка документов, которые необходимо найти и получения релевантных результатов поиска. Для этого принимают последовательность неоднозначных компонентов информации от пользователя и переводят в одну или более соответствующие последовательности менее неоднозначных компонентов информации. Эти последовательности менее неоднозначной информации предоставляются как входные данные в поисковую машину. Результаты поиска получаются от поисковой машины и представляются пользователю. Перевод между этими наборами символов и/или языками может быть выполнен посредством исследования использования терминов в выровненном тексте. Вероятности могут быть ассоциативно связаны с каждым возможным переводом. К этим вероятностям могут быть сделаны уточнения посредством исследования взаимодействий пользователя с результатами поиска. 7 н.п.ф-лы и 38 з.п.ф-лы, 16 ил.

Референт      Н.С. Лаврентьева